SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE

(ii) TITLE OF THE INVENTION: GROWTH DIFFERENTIATION FACTOR-12

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson
(B) STREET: 4225 Executive Square, Suite 1400
(C) CITY: La Jolla
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92037

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US95/
(B) FILING DATE: 12-JUL-1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Haile, Ph.D., Lisa A
(B) REGISTRATION NUMBER: 38,347
(C) REFERENCE/DOCKET NUMBER: 07265/042WO1 (FD-3830)

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 619-678-5070
(B) TELEFAX: 619-678-5099
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 12...12
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 26...26
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 29...29
(D) OTHER INFORMATION: Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCGGAATTCG GNTGGMGNVA TGGRTNRTNT AYCC 34

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base

(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCGGAATTCR CANSCRCANC YNWCNACNRY CAT 33

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGGAATTCR CANSCRCANT SNYGNACNRY CAT 33

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position

seq 51

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCGGAATTCR CANSCRCANT SNWCNACNRY CAT 33

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGGAATTCR CANSCRCANT SNBTNACNRY CAT 33

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO

sub 51

(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCGGAATTCR CANSCRCANG MNYGNACNRY CAT 33

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCGGAATTCR CANSCRCANG MNWCNACNRY CAT 33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCGGAATTCR CANSCRCANM GNYGNACNRY CAT 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCGGAATTCR CAnSCRCANM GNWCNACNMY CAT 33

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Modified Base
(B) LOCATION: 13...13
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 19...19
(D) OTHER INFORMATION: Inosine

(A) NAME/KEY:Modified Base
(B) LOCATION: 25...25
(D) OTHER INFORMATION: Inosine; Inosine also at position 28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCGGAATTCR CAnSCRCANM GNWCNACNMY CAT 33

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 360 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...357
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
CGG GCC AGG AGG AGG ACC CCC ACC TGT GAG CCT GCG ACC CCC TTA TGT      48
Arg Ala Arg Arg Arg Thr Pro Thr Cys Glu Pro Ala Thr Pro Leu Cys
 1               5                  10                  15

TGC AGG CGA GAC CAT TAC GTA GAC TTC CAG GAA CTG GGA TGG CGG GAC      96
Cys Arg Arg Asp His Tyr Val Asp Phe Gln Glu Leu Gly Trp Arg Asp
            20                  25                  30

TGG ATA CTG CAG CCC GAG GGG TAC CAG CTG AAT TAC TGC AGT GGG CAG     144
Trp Ile Leu Gln Pro Glu Gly Tyr Gln Leu Asn Tyr Cys Ser Gly Gln
        35                  40                  45
```

```
TGC CCT CCC CAC CTG GCT GGC AGC CCA GGC ATT GCT GCC TCT TTC CAT      192
Cys Pro Pro His Leu Ala Gly Ser Pro Gly Ile Ala Ala Ser Phe His
 50                  55                  60

TCT GCC GTC TTC AGC CTC CTC AAA GCC AAC AAT CCT TGG CCT GCC AGT      240
Ser Ala Val Phe Ser Leu Leu Lys Ala Asn Asn Pro Trp Pro Ala Ser
65                  70                  75                  80

ACC TCC TGT TGT GTC CCT ACT GCC CGA AGG CCC CTC TCT CTC CTC TAC      288
Thr Ser Cys Cys Val Pro Thr Ala Arg Arg Pro Leu Ser Leu Leu Tyr
                85                  90                  95

CTG GAT CAT AAT GGC AAT GTG GTC AAG ACG GAT GTG CCA GAT ATG GTG      336
Leu Asp His Asn Gly Asn Val Val Lys Thr Asp Val Pro Asp Met Val
            100                 105                 110

GTG GAG GCC TGT GGC TGC AGC TAG                                     360
Val Glu Ala Cys Gly Cys Ser
        115
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acids
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Arg Ala Arg Arg Arg Thr Pro Thr Cys Glu Pro Ala Thr Pro Leu Cys
 1               5                  10                  15

Cys Arg Arg Asp His Tyr Val Asp Phe Gln Glu Leu Gly Trp Arg Asp
            20                  25                  30

Trp Ile Leu Gln Pro Glu Gly Tyr Gln Leu Asn Tyr Cys Ser Gly Gln
        35                  40                  45

Cys Pro Pro His Leu Ala Gly Ser Pro Gly Ile Ala Ala Ser Phe His
    50                  55                  60
```

Ser Ala Val Phe Ser Leu Leu Lys Ala Asn Asn Pro Trp Pro Ala Ser
65 70 75 80

Thr Ser Cys Cys Val Pro Thr Ala Arg Arg Pro Leu Ser Leu Leu Tyr
85 90 95

Leu Asp His Asn Gly Asn Val Val Lys Thr Asp Val Pro Asp Met Val
100 105 110

Val Glu Ala Cys Gly Cys Ser
115

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2419 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 218...1267
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGCTGTGAG GTCAAGCAC AGCTATCCAT CAGATGATCT ACTTTCAGCC TTCCTGAGTC 60

CCAGACAATA GAAGACAGGT GGCTGTACCC TTGGCCAAGG GTAGGTGTGG CAGTGGTGTC 120

TGCTGTCACT GTGCCCTCAT TGGCCCCCAG CAATCAGACT CAACAGACGG AGCAACTGCC 180

ATCCGAGGCT CCTGAACCAG GGCCATTCAC CAGGAGC ATG CGG CTC CCT GAT GTC 235
Met Arg Leu Pro Asp Val
1 5

CAG CTC TGG CTG GTG CTG CTG TGG GCA CTG GTG CGA GCA CAG GGG ACA 283
Gln Leu Trp Leu Val Leu Leu Trp Ala Leu Val Arg Ala Gln Gly Thr
10 15 20

```
GGG TCT GTG TGT CCC TCC TGT GGG GGC TCC AAA CTG GCA CCC CAA GCA      331
Gly Ser Val Cys Pro Ser Cys Gly Gly Ser Lys Leu Ala Pro Gln Ala
        25                  30                  35

GAA CGA GCT CTG GTG CTG GAG CTA GCC AAG CAG CAA ATC CTG GAT GGG      379
Glu Arg Ala Leu Val Leu Glu Leu Ala Lys Gln Gln Ile Leu Asp Gly
    40                  45                  50

TTG CAC CTG ACC AGT CGT CCC AGA ATA ACT CAT CCT CCA CCC CAG GCA      427
Leu His Leu Thr Ser Arg Pro Arg Ile Thr His Pro Pro Pro Gln Ala
55                  60                  65                  70

GCG CTG ACC AGA GCC CTC CGG AGA CTA CAG CCA GGG AGT GTG GCT CCA      475
Ala Leu Thr Arg Ala Leu Arg Arg Leu Gln Pro Gly Ser Val Ala Pro
                75                  80                  85

GGG AAT GGG GAG GAG GTC ATC AGC TTT GCT ACT GTC ACA GAC TCC ACT      523
Gly Asn Gly Glu Glu Val Ile Ser Phe Ala Thr Val Thr Asp Ser Thr
            90                  95                  100

TCA GCC TAC AGC TCC CTG CTC ACT TTT CAC CTG TCC ACT CCT CGG TCC      571
Ser Ala Tyr Ser Ser Leu Leu Thr Phe His Leu Ser Thr Pro Arg Ser
        105                 110                 115

CAC CAC CTG TAC CAT GCC CGC CTG TGG CTG CAC GTG CTC CCC ACC CTT      619
His His Leu Tyr His Ala Arg Leu Trp Leu His Val Leu Pro Thr Leu
    120                 125                 130

CCT GGC ACT CTT TGC TTG AGG ATC TTC CGA TGG GGA CCA AGG AGG AGG      667
Pro Gly Thr Leu Cys Leu Arg Ile Phe Arg Trp Gly Pro Arg Arg Arg
135                 140                 145                 150

CGC CAA GGG TCC CGC ACT CTC CTG GCT GAG CAC CAC ATC ACC AAC CTG      715
Arg Gln Gly Ser Arg Thr Leu Leu Ala Glu His His Ile Thr Asn Leu
                155                 160                 165

GGC TGG CAT ACC TTA ACT CTG CCC TCT AGT GGC TTG AGG GGT GAG AAG      763
Gly Trp His Thr Leu Thr Leu Pro Ser Ser Gly Leu Arg Gly Glu Lys
            170                 175                 180

TCT GGT GTC CTG AAA CTG CAA CTA GAC TGC AGA CCC CTA GAA GGC AAC      811
Ser Gly Val Leu Lys Leu Gln Leu Asp Cys Arg Pro Leu Glu Gly Asn
        185                 190                 195

AGC ACA GTT ACT GGA CAA CCG AGG CGG CTC TTG GAC ACA GCA GGA CAC      859
Ser Thr Val Thr Gly Gln Pro Arg Arg Leu Leu Asp Thr Ala Gly His
    200                 205                 210
```

sub 51

```
CAG CAG CCC TTC CTA GAG CTT AAG ATC CGA GCC AAT GAG CCT GGA GCA      907
Gln Gln Pro Phe Leu Glu Leu Lys Ile Arg Ala Asn Glu Pro Gly Ala
215                 220                 225                 230

GGC CGG GCC AGG AGG AGG ACC CCC ACC TGT GAG CCT GCG ACC CCC TTA      955
Gly Arg Ala Arg Arg Arg Thr Pro Thr Cys Glu Pro Ala Thr Pro Leu
                235                 240                 245

TGT TGC AGG CGA GAC CAT TAC GTA GAC TTC CAG GAA CTG GGA TGG CGG      1003
Cys Cys Arg Arg Asp His Tyr Val Asp Phe Gln Glu Leu Gly Trp Arg
            250                 255                 260

GAC TGG ATA CTG CAG CCC GAG GGG TAC CAG CTG AAT TAC TGC AGT GGG      1051
Asp Trp Ile Leu Gln Pro Glu Gly Tyr Gln Leu Asn Tyr Cys Ser Gly
        265                 270                 275

CAG TGC CCT CCC CAC CTG GCT GGC AGC CCA GGC ATT GCT GCC TCT TTC      1099
Gln Cys Pro Pro His Leu Ala Gly Ser Pro Gly Ile Ala Ala Ser Phe
    280                 285                 290

CAT TCT GCC GTC TTC AGC CTC CTC AAA GCC AAC AAT CCT TGG CCT GCC      1147
His Ser Ala Val Phe Ser Leu Leu Lys Ala Asn Asn Pro Trp Pro Ala
295                 300                 305                 310

AGT ACC TCC TGT TGT GTC CCT ACT GCC CGA AGG CCC CTC TCT CTC CTC      1195
Ser Thr Ser Cys Cys Val Pro Thr Ala Arg Arg Pro Leu Ser Leu Leu
                315                 320                 325

TAC CTG GAT CAT AAT GGC AAT GTG GTC AAG ACG GAT GTG CCA GAT ATG      1243
Tyr Leu Asp His Asn Gly Asn Val Val Lys Thr Asp Val Pro Asp Met
            330                 335                 340

GTG GTG GAG GCC TGT GGC TGC AGC TAGCAAGAGG ACCTGGGGCT TTGGAGTGAA G  1298
Val Val Glu Ala Cys Gly Cys Ser
        345                 350

AGACCAAGAT GAAGTTTCCC AGGCACAGGG CATCTGTGAC TGGAGGCATC AGATTCCTGA   1358

TCCACACCCC AACCCAACAA CCACCTGGCA ATATGACTCA CTTGACCCCT ATGGGACCCA   1418

AATGGGCACT TTCTTGTCTG AGACTCTGGC TTATTCCAGG TTGGCTGATG TGTTGGGAGA   1478

TGGGTAAAGC GTTTCTTCTA AAGGGGTCTA CCCAGAAAGC ATGATTTCCT GCCCTAAGTC   1538

CTGTGAGAAG ATGTCAGGGA CTAGGGAGGG AGGGAGGGAA GGCAGAGAAA AATTACTTAG   1598

CCTCTCCCAA GATGAGAAAG TCCTCAAGTG AGGGGAGGAG GAAGCAGATA GATGGTCCAG   1658
```

CAGGCTTGAA GCAGGGTAAG CAGGCTGGCC CAGGGTAAGG GCTGTTGAGG TACCTTAAGG 1718

GAAGGTCAAG AGGGAGATGG GCAAGGCGCT GAGGGAGGAT GCTTAGGGGA CCCCCAGAAA 1778

CAGGAGTCAG GAAAATGAGG CACTAAGCCT AAGAAGTTCC CTGGTTTTTC CCAGGGGACA 1838

GGACCCACTG GGAGACAAGC ATTTATACTT TCTTTCTTCT TTTTTATTTT TTTGAGATCG 1898

AGTCTCGCTC TGTCACCAGG CTGGAGTGCA GTGACACGAT CTTGGCTCAC TGCAACCTCC 1958

GTCTCCTGGG TTCAAGTGAT TCTTCTGCCT CAGCCTCCCG AGCAGCTGGG ATTACAGGCG 2018

CCCACTAATT TTTGTATTCT TAGTAGAAAC GAGGTTTCAA CATGTTGGCC AGGATGGTCT 2078

CAATCTCTTG ACCTCTTGAT CCACCCGACT TGGCCTCCCG AAGTGATGAG ATTATAGGCG 2138

TGAGCCACCG CGCCTGGCTT ATACTTTCTT AATAAAAAGG AGAAAGAAAA TCAACAAATG 2198

TGAGTCATAA AGAAGGGTTA GGGTGATGGT CCAGAGCAAC AGTTCTTCAA GTGTACTCTG 2258

TAGGCTTCTG GGAGGTCCCT TTTCAGGGGT GTCCACAAAG TCAAAGCTAT TTTCATAATA 2318

ATACTAACAT GTTATTTGCC TTTTGAATTC TCATTATCTT AAAATTGTAT TGTGGAGTTT 2378

TCCAGAGGCC GTGTGACATG TGATTACATC ATCTTTCTGA C 2419

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 amino acids
(B) TYPE: amino acids
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Arg Leu Pro Asp Val Gln Leu Trp Leu Val Leu Leu Trp Ala Leu
1 5 10 15

Val Arg Ala Gln Gly Thr Gly Ser Val Cys Pro Ser Cys Gly Gly Ser
20 25 30

Rule 91

Lys Leu Ala Pro Gln Ala Glu Arg Ala Leu Val Leu Glu Leu Ala Lys
35 40 45

Gln Gln Ile Leu Asp Gly Leu His Leu Thr Ser Arg Pro Arg Ile Thr
50 55 60

His Pro Pro Pro Gln Ala Ala Leu Thr Arg Ala Leu Arg Arg Leu Gln
65 70 75 80

Pro Gly Ser Val Ala Pro Gly Asn Gly Glu Glu Val Ile Ser Phe Ala
85 90 95

Thr Val Thr Asp Ser Thr Ser Ala Tyr Ser Ser Leu Leu Thr Phe His
100 105 110

Leu Ser Thr Pro Arg Ser His His Leu Tyr His Ala Arg Leu Trp Leu
115 120 125

His Val Leu Pro Thr Leu Pro Gly Thr Leu Cys Leu Arg Ile Phe Arg
130 135 140

Trp Gly Pro Arg Arg Arg Arg Gln Gly Ser Arg Thr Leu Leu Ala Glu
145 150 155 160

His His Ile Thr Asn Leu Gly Trp His Thr Leu Thr Leu Pro Ser Ser
165 170 175

Gly Leu Arg Gly Glu Lys Ser Gly Val Leu Lys Leu Gln Leu Asp Cys
180 185 190

Arg Pro Leu Glu Gly Asn Ser Thr Val Thr Gly Gln Pro Arg Arg Leu
195 200 205

Leu Asp Thr Ala Gly His Gln Gln Pro Phe Leu Glu Leu Lys Ile Arg
210 215 220

Ala Asn Glu Pro Gly Ala Gly Arg Ala Arg Arg Arg Thr Pro Thr Cys
225 230 235 240

Glu Pro Ala Thr Pro Leu Cys Cys Arg Arg Asp His Tyr Val Asp Phe
245 250 255

Gln Glu Leu Gly Trp Arg Asp Trp Ile Leu Gln Pro Glu Gly Tyr Gln
260 265 270

Leu Asn Tyr Cys Ser Gly Gln Cys Pro Pro His Leu Ala Gly Ser Pro
275 280 285

Gly Ile Ala Ala Ser Phe His Ser Ala Val Phe Ser Leu Leu Lys Ala
290 295 300

Asn Asn Pro Trp Pro Ala Ser Thr Ser Cys Cys Val Pro Thr Ala Arg
305 310 315 320

Arg Pro Leu Ser Leu Leu Tyr Leu Asp His Asn Gly Asn Val Val Lys
325 330 335

Thr Asp Val Pro Asp Met Val Val Glu Ala Cys Gly Cys Ser
340 345 350